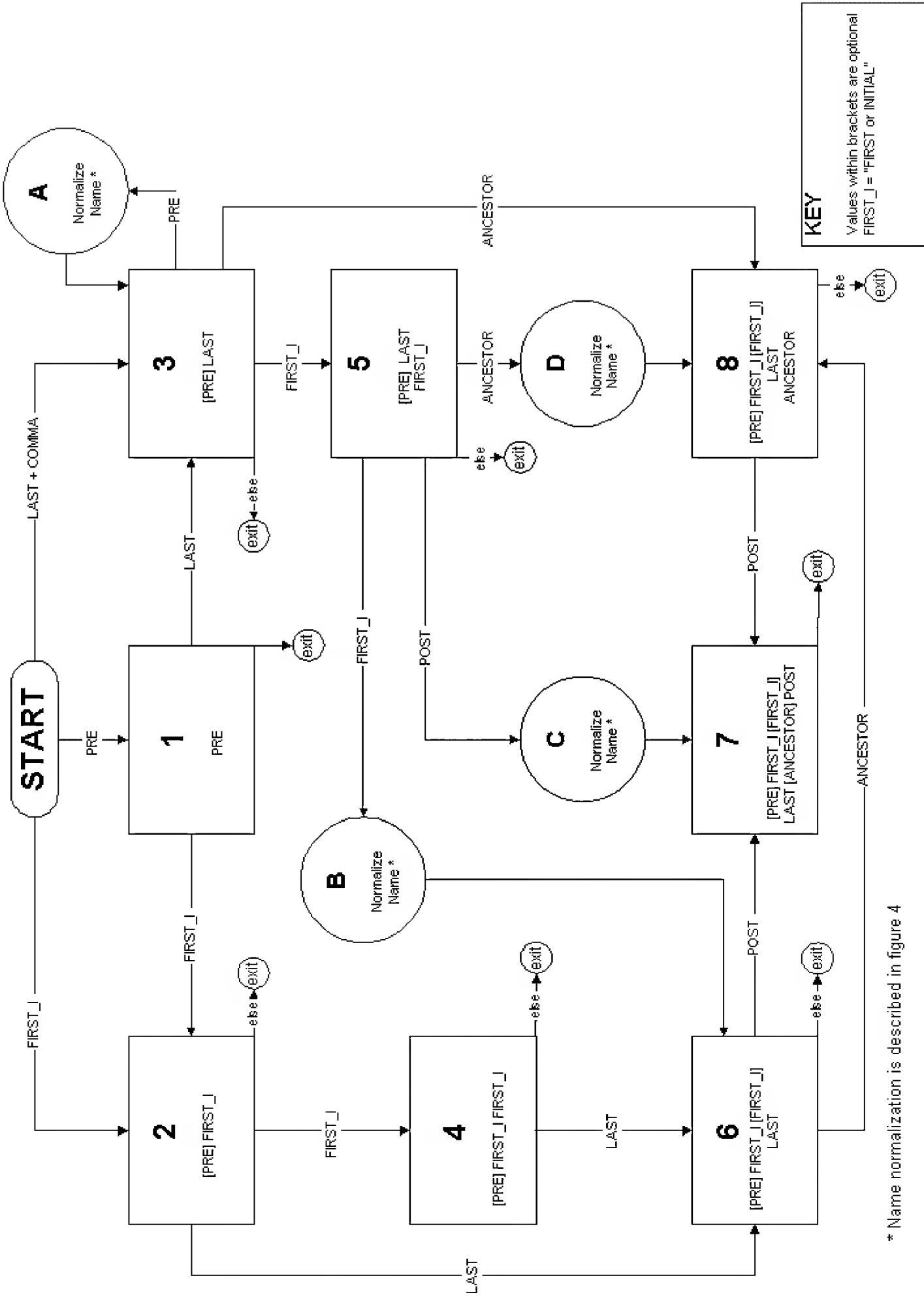


FIG. 1

First Name, Last Name Format	Extraction Algorithm States
Michael Smith	2,6
M. Smith	2,6
M. J. Smith	2,4,6
Michael J. Smith	2,4,6
M. Joseph Smith	2,4,6
Michael Joseph Smith	2,4,6
Mr. Smith	1,3
Mr. Michael Smith	1,2,6
Mr. M. Smith	1,2,6
Mr. M. J. Smith	1,2,4,6
Mr. Michael J. Smith	1,2,4,6
Mr. M. Joseph Smith	1,2,4,6
Mr. Michael Joseph Smith	1,2,4,6
Last Name, First Name Format	
Smith, Michael	3,5
Smith, M.	3,5
Smith, M. J.	3,5,7
Smith, Michael J.	3,5,7
Smith, M. Joseph	3,5,7
Smith, Michael Joseph	3,5,7
Smith, Mr.	3,3
Smith, Mr. Michael	3,3,5
Smith, Mr. M.	3,3,5
Smith, Mr. M. J.	3,3,5,7
Smith, Mr. Michael J.	3,3,5,7
Smith, Mr. M. Joseph	3,3,5,7
Smith, Mr. Michael Joseph	3,3,5,7
Post/Ancestor Combinations	
III	8
PhD	7
III PhD	8,7

Figure 2



* Name normalization is described in figure 4

Figure 3

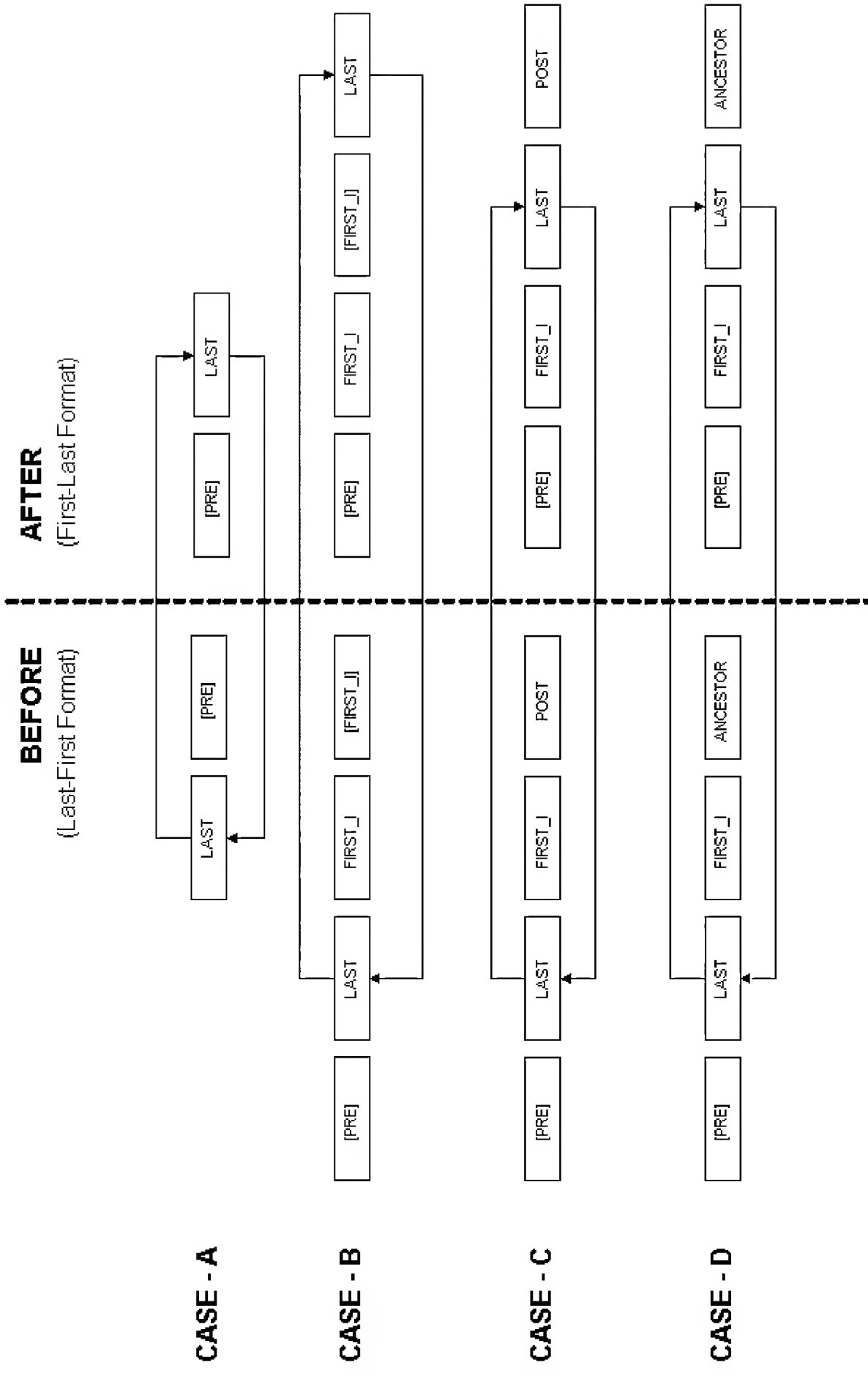


Figure 4

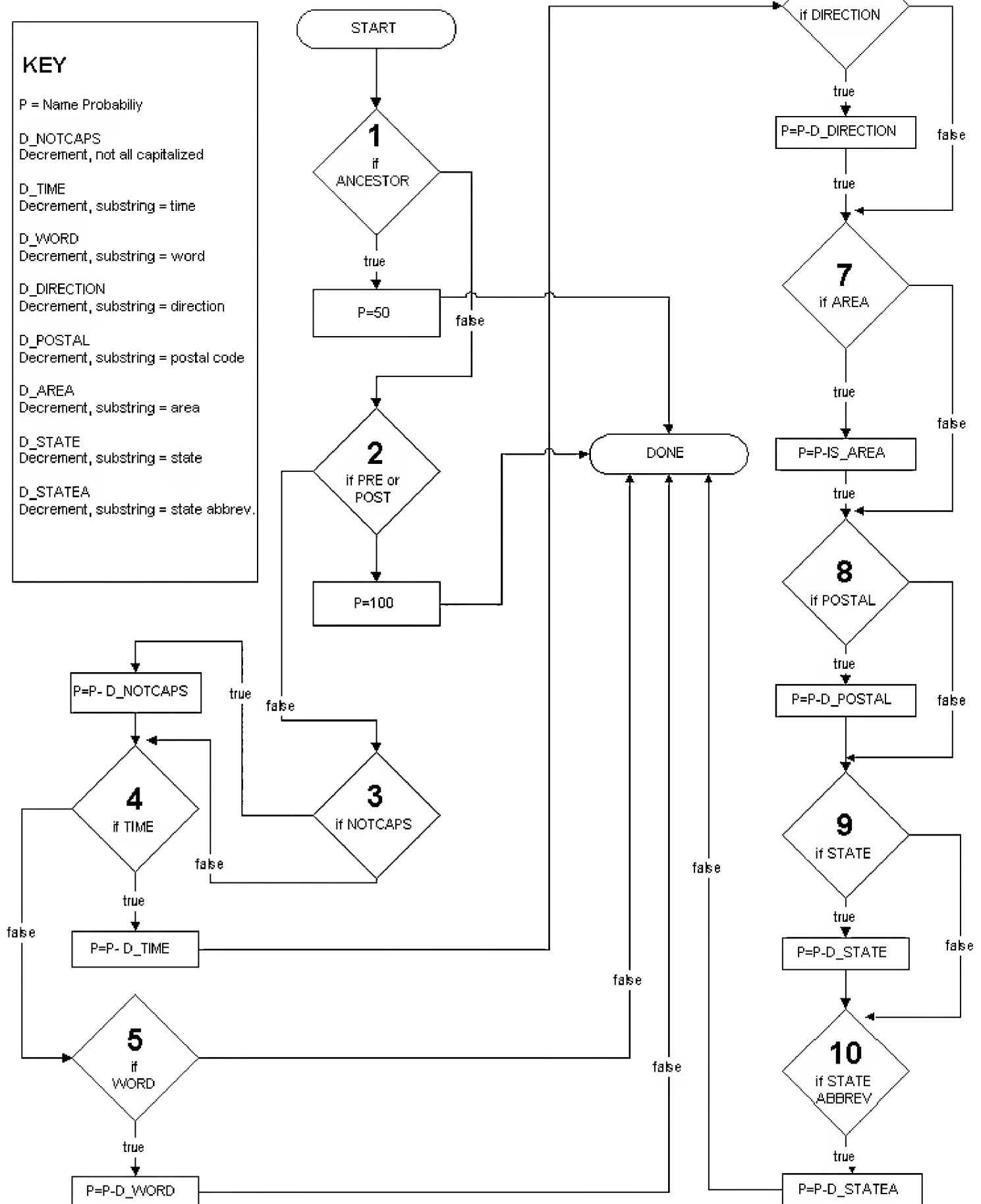
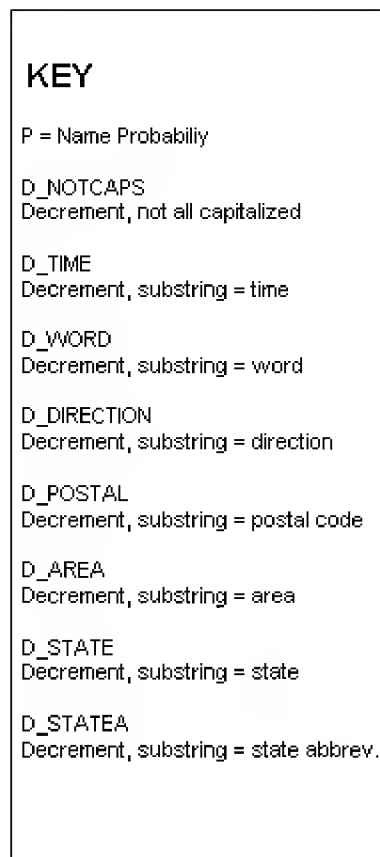


Figure 5

Processing

Network

Limits

Increments

Decrements

Coef.

Special

Freq threshold	Increment
1000	45
5000	50
10000	75
25000	60
50000	65
100000	70
200000	75
400000	80
600000	85

Add

Del

Set to default

Close

Figure 6

Processing

Network

Limits

Increments

Decrements

Coef

Special

Not cap

35

Word

45

Time

60

Direction

60

Postal

60

Areas:

State abbr.

30

State full

35

City

35

City, state

60

County

30

Country

45

Set to default

Close

Figure 7

Processing | Network | Limits | Increments | Decrements | Coef. | **Special**

Is word & small frequency threshold

Threshold for area ignore - FF and FL:

Frequency first

Frequency last

☒ 1 word & 1 word 1000

☒ 1 word 100

Name recognition threshold:

Figure 8

Processing	Network	Limits	Increments	Decrements	Coef.	Special
Pre (Mr., Dr., Mrs.)		20				
First (John, Mike)		5				
Middle (John, Mike)		2				
Last (Smith, Jones)		10				
Ancestor (Jr., III, IV)		1				
Post (Ph.D, MD, DDS)		10				
Set to default		Close				

Figure 9

$$\sum_{i=1}^6 X[i] K[i] P[i]$$

$$\sum_{i=1}^6 X[i] K[i]$$

Figure 10

1. Values for X

X[1] = PRE	X[4] = LAST
X[2] = FIRST1	X[5] = ANCESTOR
X[3] = FIRST2	X[6] = POST

2. Values for K

K[1] = PRE	K[4] = LAST
K[2] = FIRST1	K[5] = ANCESTOR
K[3] = FIRST2	K[6] = POST

3. Values for P

P[1] = PRE	P[4] = LAST
P[2] = FIRST1	P[5] = ANCESTOR
P[3] = FIRST2	P[6] = POST

Figure 11

i	Name part	Substring	X	K	P
1	PRE	Mr	1	20	100
2	FIRST1	Donato	1	5	145
3	FIRST2	S	1	2	100
4	LAST	Diorio	1	10	145
5	ANCESTOR		0	1	0
6	POST		0	10	0

Figure 12

Initial equation with values filled in from example:

$$\frac{X[1]*K[1]*P[1] + X[2]*K[2]*P[2] + X[3]*K[3]*P[3] + X[4]*K[4]*P[4] + X[5]*K[5]*P[5] + X[6]*K[6]*P[6]}{X[1]*K[1] + X[2]*K[2] + X[3]*K[3] + X[4]*K[4] + X[5]*K[5] + X[6]*K[6]}$$

Completing the equation

1.

$$\frac{1*20*100 + 1*5*145 + 1*2*100 + 1*10*145 + 0*1*0 + 0*10*0}{1*20 + 1*5 + 1*2 + 1*10 + 0*1 + 0*10}$$

2.

$$\frac{2000 + 725 + 200 + 1450 + 0 + 0}{20 + \underline{5} + 2 + 10 + 0 + 0}$$

3.

$$\frac{4375}{37}$$

4. Final name score for "Mr Donato S. Diorio" = 118.24

Figure 13

Source	Name	Context	Location
whois	Mark Tortorici	Domain Name: PROQWEST.COM Administrative Contact: Tortorici, Mark (MT3493) mark@PROQWEST.COM ProQwest 592 Weddell	
homepage	James Duran	Duran HCP Team James Duran President 408-540-0071 408-893-4905 jamesd@du	duranteam.html
homepage	Susan Fox	licon Valley recruiting today... More... Susan Fox Vice President of Operations 408-540-0072 408-	duranteam.html
homepage	Joel Abraham	ng and Life Science start ups... More... Joel Abraham Director of Midwest Operations 414-908-7799 41	duranteam.html
homepage	Jeff Polo	ograms to best meet their needs. More... Jeff Polo Client Services Manager 408-540-0073 408-893-4	duranteam.html
homepage	Bianca Medina	olunteer helping area schools... More... Bianca Medina Accounting & HR 408-540-0074 biancam@duran	duranteam.html
homepage	Peter Weddle	and employment site rankings are The Green Book by Peter Weddle , and CareerXroads, by Gerry Crispin and Mark Mehler	maynews.html
homepage	Gerry Crispin	e Green Book by Peter Weddle, and CareerXroads, by Gerry Crispin and Mark Mehler. Both do comprehensive jobs of ra	maynews.html
homepage	Mark Mehler	ter Weddle, and CareerXroads, by Gerry Crispin and Mark Mehler . Both do comprehensive jobs of ranking and classi	maynews.html

Figure 14